Journal of Machine Learning Research 1 (2000) 1-48 Submitted 4/00; Published 10/00

Accelerated oblique random survival forests

Marina Meilă MMP@STAT.WASHINGTON.EDU
Department of Statistics
University of Washington
Seattle, WA 98195-4322, USA

Michael I. Jordan JORDAN@CS.BERKELEY.EDU
Division of Computer Science and Department of Statistics
University of California
Berkeley, CA 94720-1776, USA

Editor: Kevin Murphy and Bernhard Schölkopf

Abstract

abstract here

Keywords: Bayesian networks, mixture models, Chow-Liu trees

1. Introduction

2. Related work

Table 1 is an example of a referenced LaTeX. Several machine learning algorithms can engage with right-censored time-to-event outcomes. In the current study, we consider four classes of learners: random forests, boosting ensembles, regression models, and neural networks.

accelerated oblique random survival forests `aorsf`
original oblique random survival forests `obliqueRSF`
axis-based random survival forests `randomForestSRC & ranger`
axis-based conditional inference forests `party`
gradient boosted decision trees `xgboost`

3. Results

	Performance metric (SD)		Computing time, seconds			
			Fit model		Predict risk	
	C-Statistic	Scaled Brier	Mean	Ratio	Mean	Ratio
actg						
Party	0.772 (0.058)	0.076 (0.042)	2.181	2.19	4.889	4.62
aORSF(i=1)	0.755 (0.059)	0.057 (0.055)	0.998	1.00	1.057	1.00
Ranger	0.754 (0.064)	0.064 (0.027)	0.397	0.398	0.510	0.483
aORSF(i=15)	0.754 (0.063)	0.052 (0.056)	2.498	2.50	2.559	2.42
Rfsrc	0.752 (0.064)	0.061 (0.047)	0.775	0.777	0.833	0.787

(continued)

	C-Statistic	Scaled Brier	Mean	Ratio	Mean	Ratio
aorsf_net	0.750 (0.059)	0.055 (0.064)	28.449	28.5	28.510	27.0
breast						
aORSF(i=15)	0.764 (0.008)	0.143 (0.016)	11.867	2.56	12.403	2.43
Ranger	0.756 (0.013)	0.102 (0.017)	0.555	0.120	0.724	0.142
Party	0.753 (0.013)	0.118 (0.016)	8.030	1.73	10.445	2.05
aORSF(i=1)	0.751 (0.002)	0.120 (0.012)	4.636	1.00	5.106	1.00
aorsf_net	0.741 (0.022)	0.062 (0.049)	585.625	126.3	585.887	114.7
Rfsrc	0.707 (0.005)	0.049 (0.039)	1.354	0.292	1.522	0.298
colon						
Rfsrc	0.815 (0.007)	0.279 (0.007)	1.854	0.581	2.026	0.600
aORSF(i=15)	0.778 (0.007)	0.199 (0.018)	7.515	2.36	7.709	2.28
aORSF(i=1)	0.776 (0.004)	0.201 (0.014)	3.190	1.00	3.378	1.00
aorsf_net	0.753 (0.006)	0.156 (0.018)	149.637	46.9	149.809	44.3
Party	0.753 (0.012)	0.183 (0.008)	3.968	1.24	10.451	3.09
Ranger	0.738 (0.008)	0.147 (0.005)	1.795	0.563	2.398	0.710
gbsg2						
Ranger	0.704 (0.076)	0.090 (0.045)	0.328	0.339	0.450	0.438
Party	0.686 (0.042)	0.092 (0.048)	0.746	0.773	2.055	2.00
aORSF(i=15)	0.685 (0.048)	0.077 (0.066)	2.077	2.15	2.144	2.09
Rfsrc	0.676 (0.032)	0.067 (0.052)	0.931	0.963	1.014	0.986
aORSF(i=1)	0.674 (0.045)	0.062 (0.062)	0.966	1.00	1.028	1.00
aorsf_net	0.670 (0.038)	0.068 (0.056)	57.024	59.0	57.087	55.5
guide_it						
aORSF(i=15)	0.730 (0.009)	0.133 (0.008)	4.705	2.67	4.811	2.57
aORSF(i=1)	0.728 (0.014)	0.132 (0.012)	1.762	1.00	1.869	1.00
aorsf_net	0.728 (0.020)	0.133 (0.015)	86.476	49.1	86.590	46.3
Party	0.705 (0.021)	0.101 (0.010)	1.800	1.02	3.727	1.99
Rfsrc	0.705 (0.032)	0.109 (0.027)	1.000	0.568	1.104	0.591
Ranger	0.704 (0.016)	0.102 (0.001)	0.663	0.376	0.855	0.457
Mayo Clinic Primary Biliary Cholangitis Data, N = 276						
aORSF(i=1)	0.787 (0.094)	0.202 (0.188)	2.204	1.00	2.400	1.00
aORSF(i=15)	0.787 (0.100)	0.203 (0.186)	5.038	2.29	5.242	2.18
Ranger	0.783 (0.091)	0.161 (0.138)	0.709	0.322	1.053	0.438
aorsf_net	0.782 (0.101)	0.190 (0.202)	128.300	58.2	128.469	53.5
Party	0.779 (0.094)	0.185 (0.167)	3.291	1.49	6.998	2.92
Rfsrc	0.778 (0.113)	0.195 (0.187)	0.956	0.434	1.076	0.448
Overall						
aORSF(i=1)	0.946 (0.018)	0.493 (0.083)	0.344	1.00	0.376	1.00
aorsf_net	0.939 (0.018)	0.484 (0.052)	23.093	67.2	23.123	61.6
Party	0.936 (0.022)	0.454 (0.022)	0.288	0.839	0.518	1.38

(continued)

	C-Statistic	Scaled Brier	Mean	Ratio	Mean	Ratio
aORSF(i=15)	0.936 (0.014)	0.474 (0.080)	0.744	2.17	0.773	2.06
Rfsrc	0.927 (0.012)	0.462 (0.070)	0.363	1.06	0.397	1.06
Ranger	0.926 (0.014)	0.344 (0.055)	0.044	0.127	0.085	0.226
Rotterdam Breast Cancer Data, N = 2,982						
Ranger	0.719 (0.011)	0.145 (0.011)	2.480	0.410	4.456	0.692
aorsf_net	0.714 (0.007)	0.141 (0.007)	237.007	39.2	237.337	36.9
Party	0.712 (0.008)	0.141 (0.012)	7.030	1.16	21.888	3.40
aORSF(i=1)	0.712 (0.006)	0.139 (0.008)	6.053	1.00	6.436	1.00
aORSF(i=15)	0.711 (0.005)	0.139 (0.007)	13.541	2.24	13.925	2.16
Rfsrc	0.701 (0.008)	0.120 (0.010)	2.222	0.367	2.458	0.382
time_to_million						
aorsf_net	0.937 (0.013)	0.560 (0.040)	78.384	55.8	78.461	53.0
aORSF(i=15)	0.937 (0.018)	0.555 (0.043)	4.491	3.20	4.569	3.09
aORSF(i=1)	0.935 (0.017)	0.552 (0.039)	1.405	1.00	1.481	1.00
Rfsrc	0.932 (0.023)	0.550 (0.045)	0.784	0.558	0.853	0.576
Party	0.921 (0.014)	0.485 (0.034)	0.621	0.442	1.684	1.14
Ranger	0.916 (0.023)	0.452 (0.039)	0.324	0.231	0.428	0.289
vdv						
Rfsrc	0.782 (0.231)	0.033 (0.185)	0.198	0.080	0.450	0.148
aorsf_net	0.770 (0.161)	-0.019 (0.283)	20.190	8.13	20.749	6.80
aORSF(i=1)	0.762 (0.079)	-0.012 (0.126)	2.484	1.00	3.050	1.00
Ranger	0.751 (0.070)	-0.017 (0.093)	0.484	0.195	0.576	0.189
aORSF(i=15)	0.735 (0.171)	-0.025 (0.135)	2.569	1.03	3.136	1.03
Party	0.722 (0.102)	-0.030 (0.165)	8.098	3.26	14.088	4.62
veteran						
Ranger	0.865 (0.063)	0.183 (0.036)	0.023	0.113	0.043	0.194
aORSF(i=15)	0.837 (0.048)	0.281 (0.075)	0.371	1.84	0.387	1.75
aORSF(i=1)	0.833 (0.053)	0.278 (0.081)	0.202	1.00	0.222	1.00
Party	0.833 (0.057)	0.235 (0.056)	0.150	0.744	0.231	1.04
aorsf_net	0.820 (0.059)	0.256 (0.083)	17.115	84.7	17.133	77.3
Rfsrc	0.784 (0.111)	0.218 (0.112)	0.077	0.380	0.099	0.447

Acknowledgments

We would like to acknowledge support for this project from the National Science Foundation (NSF grant IIS-9988642) and the Multidisciplinary Research Program of the Department of Defense (MURI N00014-00-1-0637).

Learner	Software	Description
Random Survival Forests		
Standard	`RandomForestSRC`	Axis based survival trees following Leo Breiman's original random forest algorithm, with cut-points selected to maximize a log-rank statistic.
Oblique	`obliqueRSF` `aorsf`	Oblique survival trees following Leo Breiman's random forest algorithm but not using random coefficients for linear combinations. `obliqueRSF` uses penalized models, and `aorsf` additionally uses partially trained models. Cut-points are selected to maximize a log-rank statistic.
Extremely Randomized	`ranger`	Axis-based survival trees grown with randomly selected features and cut-points
Conditional Inference	`party`	Axis based survival trees grown using unbiased recursive partitioning.
Boosting ensembles		
Trees	`xgboost`	The Cox likelihood function is maximized additively with decision trees. Nested cross validation (5 folds) is applied to tune the number of trees grown.
Models	`xgboost`	The accelerated failure time likelihood function is maximized additively with decision trees. Nested cross validation (5 folds) is applied to tune the number of trees grown.
Regression models		
Cox Net	`glmnet`	The Cox model is fit using an elastic net penalty. Nested cross validation (5 folds) is applied to tune penalty terms.
Neural networks		
Cox Time	6	87837

Table 1: Table to test captions and labels.

Appendix A.

In this appendix we prove the following theorem from Section 6.2:

Theorem *Let u, v, w be discrete variables such that v, w do not co-occur with u (i.e., $u \neq 0 \Rightarrow v = w = 0$ in a given dataset $\mathcal{D}$). Let N_{v0}, N_{w0} be the number of data points for which $v = 0, w = 0$ respectively, and let I_{uv}, I_{uw} be the respective empirical mutual information* *values based on the sample $\mathcal{D}$. Then*

$$N_{v0} \;>\; N_{w0} \;\;\Rightarrow\;\; I_{uv} \;\leq\; I_{uw}$$

with equality only if u is identically 0. ■

Proof. We use the notation:

$$P_v(i) \;=\; \frac{N_v^i}{N}, \quad i \neq 0; \quad P_{v0} \;\equiv\; P_v(0) \;=\; 1 - \sum_{i \neq 0} P_v(i).$$

These values represent the (empirical) probabilities of v taking value $i \neq 0$ and 0 respectively. Entropies will be denoted by H. We aim to show that $\frac{\partial I_{uv}}{\partial P_{v0}} < 0$....

Remainder omitted in this sample. See http://www.jmlr.org/papers/ for full paper.